DOCUMENT RESUME

ED 453 291 TM 032 830

AUTHOR	Wilkerson, LuAnn; Rose, Mike
TITLE	Learning from the Narrative Comments of Standardized Patients during an Objective Structured Clinical Examination of Fourth-Year Medical Students.
PUB DATE	2001-00-00
NOTE	13p.
PUB TYPE	Reports - Research (143)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Clinical Experience; *Empathy; Higher Education; Medical Education; *Medical Students; *Personal Narratives; Qualitative Research
IDENTIFIERS	Caring; *Standardized Patients

ABSTRACT

The standardized patient (SP) examination is used in a majority of medical schools to test clinical skills. This exam usually yields both numerical ratings of clinical skill and narrative comments by patients or observers, yet most empirical studies of SP assessment focus on the numerical ratings only. This study analyzes the comments on a recent administration of the examination qualitatively. The comments of 16 SPs portraying 8 clinical cases give a fuller sense of the meaning of general descriptors like "empathy" and "care." The comments also provide a window into the well-documented variability in the SP communication skills domain, suggesting the domain-specificity of such skills. (Contains 2 tables and 14 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 453 291

Learning from the Narrative Comments of Standardized Patients During an Objective Structured Clinical Examination of Fourth-Year Medical Students

LuAnn Wilkerson and Mike Rose
UCLA

TM032830

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

L. Wilkerson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

Learning from the Narrative Comments of Standardized Patients During an Objective Structured Clinical Examination of Fourth-Year Medical Students

LuAnn Wilkerson and Mike Rose
UCLA

Abstract:
The standardized patient (SP) examination is used in a majority of medical schools to test clinical skills. This exam usually yields both numerical ratings of clinical skill and narrative comments by patients or observers, yet most empirical studies of SP assessment focus on the numerical ratings only. The present study qualitatively analyzes the comments on a recent administration of the exam. The comments give a fuller sense of the meaning of general descriptors like "empathy" and "care". As well, the comments provide a window onto the well-documented variability in the SP communication skills domain, suggesting the domain-specificity of such skills.

Introduction:

The standardized patient examination has become a fixture in the majority of medical schools. In 1999, 73.5 percent of the senior medical students responding (80%) reported that an objective structured clinical examination (OSCE) involving standardized patients (SPs) was used in evaluating their clinical skills (AAMC, 1999). There is a rapidly growing literature on the utility and reliability of such examinations for the evaluation of both medical students and residents (Colliver & Swartz, 1997; Holmboe & Hawkins, 1998). However, all of these studies focus on the numerical ratings of clinical skills provided by the patients or observers. What do the narrative comments provided by the standardized patients tell us about students' performance? What are the dimensions of performance associated most frequently with the SP's satisfaction with an encounter and do these dimensions differ among cases? Do the narratives provide a reliable measure of a student's performance across cases? The measurement literature on OSCEs reports that reliability using checklist scores is moderate, .50-.60 (Vu & Barrows, 1994), supporting the concept of domain-specificity in competency. Is this same variability represented in the narrative comments that students receive across an entire exam?

Method

We conducted a qualitative analysis of narrative comments written by 16 SPs portraying eight clinical cases during a clinical performance examination for senior medical students at UCLA School of Medicine in the summer of 1998. The multiple-station exercise consisted of eight standardized patient encounters in which students were instructed to perform focused histories and/or physical examinations with attention to skills in patient/physician interaction and counseling. The eight cases were developed by a consortium of faculty members from five medical schools to represent a mix of acute, chronic, well-care, behavioral, grave prognosis, and ill-defined presentations. Since students take the half-day examination in small groups over a three week period, we selected responses from six half days for analysis with two sessions taken from early in

the exam period, two from the middle, and two from the final period. During these sessions, 45 senior medical students were evaluated by the SPs. This sample represents 25% of the students tested.

The SPs wrote narrative comments in response to two questions at the end of each scoring checklist. In the first, the SP was asked to indicate satisfaction with the encounter in a simple yes/no question and to write comments regarding "satisfaction with this student physician encounter." A second question asked for comments from the patient on aspects of the patient/physician interaction items taken from the Calgary-Cambridge Observation Guide (Kurtz and Silverman, 1996).

We began our analysis by reading through the entire set of comments for both questions and tagged central features. By writing analytic memos to one another, we recorded our reflections, biases, and identified emerging themes in the narratives (Miles and Huberman, 1984). We continued our discussion of the emerging themes until we reached agreement on a set of codes to be used in categorizing segmented phrases in each of the narrative statements. Given the tendency of the SP to focus satisfaction comments on aspects of communication, we combined the two sets of responses for coding.

Finally, we considered the possible relationships among the themes by examining the narrative comments from three perspectives. First, we determined themes across all eight cases. Second, we identified themes for each of the eight cases to determine if there were differences among the cases in the themes cited as strengths or weaknesses. Third, we considered comments for each student across the eight cases to address the question of reliability. In order to examine this issue of consistency, we coded each patient's comments about an individual student as completely positive, mixed positive and negative, or totally negative. In determining negativity, we chose to eliminate those comments that suggested test anxiety that was overcome during the encounter. Finally, we assigned an overall score to each student using the following scale:

1 = consistently positive: only one or two patients' comments were negative or mixed.

2 = variable: three or four patients' comments were negative or mixed

3 = consistently negative: five or more patients' comments were negative or mixed.

One of the authors (LW) scored all 45 cases. The other author (MR) scored 23 cases for comparative purposes. We discussed differences until we reached 100 % agreement. We subsequently interviewed two of the students who fell in category three in an attempt to better understand the source of the negative comments. Finally, we examined the validity of this scoring system by comparing the narrative score assigned for each student to the overall score for patient/physician interaction resulting from the checklist completed by the SPs.

Results

All but two of the 16 SPs wrote comments on the checklist resulting in a total of 345 narratives. Two of the SPs provided comments on some students and not on others. In segmenting the narratives into discrete descriptive events, we identified 593 segments to

which we assigned a code. This resulted in an average of 13.2 coded segments per student.

Across Cases. Four themes emerged in the analysis of narratives across all of the cases. Most of the comments, both positive and negative, referred to issues of communication and interpersonal skills, focusing on either (1) making a connection with the patient or (2) putting the patient at ease. However, in 80 of the encounters, the SPs also commented on issues related to (3) fund of knowledge, particularly in relation to the explanations provided by the student, or (4) professional behaviors.

Weaknesses were both in the affective and behavioral areas. An average of 8 students were cited for weaknesses in each case, with the two very medical cases (abdominal pain, cough) having the fewest negative comments and the two heavily psychosocial cases (adolescent visit, new mother), having the most. Weak students were described by the SPs as not connecting, not caring, being disinterested, or ignoring emotional needs of the patient.

Making a Connection. The most frequent comments (83) by the SPs concerned the students' skills in making a personal connection with the patient. The patients rarely used the term "empathic" in describing this connection, using instead the terms caring, concerned, humane. They described these students as being caring and concerned or condescending, arrogant, or detached. Positive behaviors included smiling appropriately, listening, encouraging discussion, and directly stating concern. Negative behaviors were talking too much, not listening, using jargon, not making eye contact, placing oneself at a distance from the patient.

> "The doctor made me feel very well cared for. He was reassuring and allayed my fears. I would definitely come back to him. He answered all my questions. I liked the way he was concerned that I might have an emotional response. I felt that he would support me that way, also."
>
> "I felt like she was on my side and was trying to find a way for me to deal with my condition."
>
> "As soon as he came into the room and sat down, he scooted the chair as far away from me as possible. Eye contact was wandering, very little warmth or concern. He offered no assistance (though he knew how much pain I was in)."
>
> "Borderline encounter: she looked like a doctor, she questioned me smoothly and efficiently, and she started off with caring eyes. But my gut said no. She left me dangling, without any options, without hope. As the encounter (quickly) progressed, she looked at me as a 'case' rather than a person. Treated me like just another patient with a disease. Never talked to me."

Putting the Patient at Ease. The second most frequently cited strength was the ability to make the patient feel comfortable, and at ease, both emotionally and physically (52).

Patients described the student as calming, comforting, reassuring or cold, scary, unsure, nervous, or aggressive. Positive behaviors included not rushing, allowing questions, making supportive statements, conducting the physical exam to minimize discomfort, and directly naming the patient's anxiety. In the chest pain case, for example, 39 of the 45 students in the study were cited for easing the patient's anxiety or reducing her fears. Negative behaviors described were rushing, being rough during the physical exam, ignoring the patient's emotional messages, mumbling, or interrupting.

> "It was so great how she started off by telling me that she wanted to talk to me without my parents so that she could answer any questions about stuff I didn't want them to know. She kept telling me that I was special and important…She has a great smile and made me feel like a new person when it was over! A++++++ It doesn't get any better than this."

> "This doctor made me feel very well cared for. He was reassuring and allayed my fears. I would definitely come back to him. He answered all my questions. I liked the way he was concerned that I might have an emotional response. I felt that he would support me."

> "This doctor scared me. He made me feel like an object. Also, he did not attend to my emotional needs. His alternatives for treatment seemed very extreme. He did not respond to my tears. At times he was condescending, like when he said he would look into psychiatric help should I have cancer."

Using Knowledge to Help the Patient. Students were also praised for being knowledgeable, particularly for the ability to use their knowledge in providing clear explanations and conducting the encounter in an organized way. Such students were described as helpful, interested, knowledgeable. Positive behaviors included using common terms, giving an appropriate amount of information for the patient, and providing clear explanations.

> "I loved the way he explained the reason for things during the physical."

> She was well informed and seemed confident in her abilities and answers. When I didn't understand something (i.e., shot terminology), she checked with me and explained in plain English what she was talking about.

> "Very knowledgeable (and shared all sorts of new information with me). Very informative. Terrific explanations. I enjoyed our time together."

The negatively described students tended to talk too much and listen too little, creating a one-way communication event. They did not elicit the patient's questions nor check for understanding. The answers that they provided were too complex or jargon-laden to be useful to the patient. Some were described as stating outcomes without regard for their impact. Patients labeled them as arrogant, unsure, or disinterested.

"As the encounter progressed, he became more and more lost in thought, like I was a puzzle he wanted to solve rather than a person. During the physical exam he paused to think by staring right at me, clicking his tongue…stared at his clipboard for long stretches, like he forgot I was there."

"He overloaded me with having surgery. He did not ask if I had any questions or concerns. Here was no two-way transmission."

Acting Professionally. Students were praised by the SPs for seeming competent, being thorough, acting appropriately confident, and appearing nonjudgmental. Such students had an organized approach to the visit, not needlessly repeating themselves, completing the encounter in the time allotted, and projecting nonverbal and verbal messages that were consistent.

"Her questions were almost conversational, not 'doctory.' Easy, relaxed interview style. Questions never came out of nowhere, they always flowed one to the next."

"End speech lovely: sat back, thoughtfully explained things, took his time, seemed authoritative, yet caring. Never rushed."

There were far more negative than positive comments in this category than positive ones with SPs directly labeling students as unprofessional because they seemed insincere, condescending, rushed, or scattered. Of particular concern were those students who seemed not to take the visit seriously, laughing at inappropriate times, avoiding eye contact, repeating questions that the patient had just answered, or leaving early.

"He didn't respect me. He laughed. He overloaded me. He left me on the table while he checked the folder 4 times. The use of his voice was a mimicry, joke kind of voice. He didn't empathize when I asked for pain meds. He laughed then said uhhh…I want a second opinion and his license revoked!!!"

"He mumbled and looked away quite a bit. He used med speak several times and sometimes I couldn't hear. He went almost directly to TB."

"I had lots to tell him; he seemed not to want to hear it and would move on to his next question. There was no swaying this jury -- the defendant was guilty, guilty, guilty."

"He kept asking me the same questions which annoyed me, I thought he didn't believe me. Example. Did you get in trouble with your friends before? No. I never get to hang out with them,. Two minutes later same question. Same answer then "When you do hang out what do you do? I don't' get to hang out with them! Why don't you listen to me?!"

By Case. Though some communication skills emerged as important across cases, the narratives within each case clustered around some themes and not others. The cases differed significantly in the challenges that they posed to the student. Differences

included features of the problem, content domain involved, complexity of the complaint, number of tasks required, degree of psychosocial intensity, patient age, patient sex, patient personality. This variability between cases was reflected in a unique satisfaction or interact ional profile for each case, e.g., what seems to matter to this patient. For example, in one case, the student had to give bad news to a patient who reacts with emotional distress. This exchange provided the most complex interact ional challenge for the students since they had to combine knowledge of disease management and prognosis with an ability to provide the emotional support needed when giving bad news (Using Knowledge, Professionalism). A very different case was that of a patient with acute low back pain where attentive listening and caution in conducting the movements of the physical exam were the communication strengths most frequently mentioned by the patients (Putting the Patient at Ease). The noncompliant diabetic commented most frequently on the students' ability to make him feel comfortable and available for questions (Connecting with the Patient). For the new mother, the ability to provide clear, thorough, helpful explanations were cited more frequently than other characteristics (Using Knowledge).

Consistency Across Cases. Using a rating scale to indicate the positive and negative quality of the standardized patient comments, we examined the narratives for each individual student across all the cases. The need for an organizational coding system developed after we began to notice that a subset of the students had a great deal of variability in how they were described. For example, one student was sequentially described as:
More interested in the exam than in the patient
Concerned and good at listening
Felt like she cared
Information was good but occasionally confusing
Eased my fears.
Made me feel like a little kid, patronizing.
Felt like she didn't believe me.

Twenty-one of the students (47%) were consistently cited by the various patients for positive aspects of communication. For these students, only one or two of the eight patients provided any negative comment. Two students received no negative comments from any patient. They were both described as caring, personable, considerate, articulate, comforting, and confident.

In order to validate our system of rating the narrative comments, we examined, using an ANOVA, the relationship between our three categories, consistently positive, variable, and consistently negative, and the total percent correct for the Physician/Patient Interaction items across all eight cases based on SP checklist scores. The Consistently Positive group (M=93%) and the Variable group (M=88%) received significantly higher total percent correct scores ($F=13.73$, $df=2, 42$, $p<.0001$) than did the Consistently Negative group (M=80%). There was not a significant difference between the means of the Consistently Positive and Variable groups.

Mean Total % Correct for Physician/Patient Interaction Checklist Items for Each Narrative Rating Group

Narrative Rating	N	Mean % (SD) Correct PPI
Consistently positive (1)	21	93 (3.9)*
Variable (2)	13	88 (7.0)*
Consistently negative (3)	11	80 (8.7)

*p<.0001 with consistently negative

Thirteen students (29%) received mixed positive and negative comments or only negative comments from three or four patients. Eleven students (24%) received mixed or negative comments from five or more patients. Different patients often commented on the same communication behavior in these more negatively cited students. For example, one of these students was described by five patients as being unsure, overly tentative, or not confident. Seven out of the eight patients described one student as mumbling. Another student was described as aggressive or arrogant by four different patients.

Interviews with two of these students indicated that they were largely unaware of the way in which they were affecting the patient, e.g., that mumbling may be interpreted as a lack of confidence, or that stopping to think may be perceived as demonstrating a lack of interest in the patient. For example, a patient described one of these students as running "through a mental checklist" rather than really connecting to her. In the subsequent interview, the student described that process as his strategy for pursuing a diagnosis:

> "So her chief complaint is cough. And naturally cough can be a myriad of things. Cough can be related to any of the anatomical structures in that area. The thorax, the lungs, heart, the gastrointestinal tract, anything in your head or neck. So I need to better delineate what organ structure we're talking about. So naturally I want to know about how long has she had the cough, chronicity of the cough that could imply something like an acute infection versus like a lung process like cancer. And I want to know about associated symptoms, is she coughing anything up? If there is like a lot of bloody sputum, which she does have, that is indicative of something in her lungs. That is more indicative of something more severe. Something enough to cause erosion or rupture to any vessel will cause blood. It it's just like um you know there is nothing, like a little sore throat, that's probably a little less acute to her situation. So you want to gauge how sick is this person. That's a very important criterion. Um, other things to know are, you know, associated symptoms…In terms of symptomology you want to know if she has like fatigue, weight loss, a variety of situations can cause systemic complaints which can often point to more serious maladies. Like if you have cancer or if you have tuberculosis, you have HIV infection, you often can get fever, fatigue, and night sweats. So the better to figure out other kinds of structures I asked her like a bunch of screening questions. Like did she have any problems with her head, ears, eyes, throat. She had nothing. Did she have any kind of chest pain, anything indicative

of heart problems, any kind of exceptional pain? In her age group, I mean that's very rare but something to screen for…So I'm not thinking about the heart and I'm not thinking it's her head. So I mean the money is with her lungs. It's probably not her GI tract, you know, we asked about problems with you know, not really swelling but any problems with eating, any nausea, vomiting. None of that was positive so I just focused on the lungs."

Discussion

It is not surprising that the SPs narrative comments most frequently referred to the interpersonal and communication skills of the students. These skills have been associated with standardized patient satisfaction (Blue et al, 2000) and were stimulated by the question of satisfaction with this encounter. It is interesting that 80 of the 345 narratives also noted issues related to the student's knowledge, particularly as that knowledge was represented in clear, thorough, but concise explanations of diagnostic, management, or preventive concerns. Colliver et al (1999) found a moderate correlation between scores for history taking and physical examination, defined by the authors as the cognitive dimension of clinical competence and scores for interpersonal skills. They concluded that these two dimensions are interdependent, with "each affecting and being affected by the other" (pg. 273).

Empathy has been called the most important characteristic of a good physician (Spiro et al 1993.) The narrative comments of the SPs in this study bore out the importance of this skill, although not the term itself. More frequently than any other aspect of the interaction, the SPs commented on the students' ability to connect with them as persons. Behaviors associated with this connection included appropriate eye contact, attentive listening, direct expressions of concern, and care in conducting the physical examination to avoid causing pain. In a study of a single checklist item on empathy in an OSCE, Colliver et al, (1998) were able to document empathic behavior on the part of students in the majority of cases but were unable to provide descriptions of this behavior such as those revealed in the narrative comments of the SPs in our study. Specific descriptions provide a tool for teaching aspects of empathy to students or to SPs who are being asked to assess them.

The finding of variability of narrative themes across cases and individual variation among a single student raises the issue of domain specificity in communication skills, a concept already well recognized in the cognitive domain (Vu & Barrows, 1994). A case for the domain specificity of communication skills could be made if two conditions were satisfied: (1) variability in communication themes were found to exist among patient stations and (2) this variability had some relationship to the content of the station.

Consistency Across Cases in Communication Skills: Existing Studies of the Objective Structured Clinical Examination

Studies	Variability among stations	Links with Content of the Station
Hodges 1996	X	X
Colliver 1998	X	X
Colliver 1999	X	X
Donnelly 2000	X	
Wilkerson 2001	X	X

Hodges et al (1996) identified the domain-specific nature of communication skills in a study of the reliability of OSCE stations specifically designed to measure communication skills. They concluded that "'communication skills' are highly bound to content and that increased difficulty and increased score variance alone are not enough to improve generalizability." (p. 42). The concept of a case-specific dimension of communication is further supported in two recent studies. Colliver et al (1998) found an overall reliability of .43 across seven OSCE cases for a single empathy item. History taking and physical examination scores were significantly lower for those students who were scored as empathic on fewer than half the cases. In a comparison of the ratings of faculty and SPs on interpersonal skills, Donnelly et al (2000) found statistically significant differences in the mean score for each OSCE problem with between rater correlations of .6 but between problem correlations for a single resident of only .20. "It may be that interpersonal skills, like clinical reasoning skills, are affected by the context of the clinical case" (pg. S95). In our own study, the best surrogate for knowledge might be the history since completing items on the checklist requires a certain knowledge of the diagnosis and what history and physical examination items would be essential.

History as a Measure of Content in the CPX

Narrative Rating	N	Mean % (SD) Correct HX
1. Consistently positive	21	69.1 (5.8)*
2. Variable	13	69.5 (5.2)*
3. Consistently negative	11	61.0 (12.8)

*$p<.05$ with consistently negative

We are in agreement with Hodges et al (1996) that a generalized communication skills checklist for an OSCE station may be less reliable than one built around specific aspects of communication needed for that particular case.

If communication skills are domain-specific, it is not surprising that we found a lack of consistency in the comments about an individual student across the standardized patients.

Although half of the students in our study received fairly consistent positive or negative narrative evaluations across cases, one-third received decidedly mixed evaluations. One patient would describe a student as "wasn't listening" while another praised her for "great communication." However, domain specificity aside, is it reasonable to expect beginning fourth year medical students to demonstrate consistent communication skills across a variety of patient problems? Within our sample of fourth-year medical students, there was variation in interests and areas of specialization and some, though not profound, variability in training that could account for differences in performance across cases. There is the possibility suggested by Hodges et al (1996) that a student's assurance about his or her own knowledge could affect communication skill. In addition, there are possible psychodynamic, sociological, and cultural factors that can affect performance, as well as issues related to the dynamics of professional role development (Rose & Wilkerson, In press). Given the multi-dimensional nature of the clinical encounter, we wonder if some variation in scores in a series of such complex human interactions might be expected at this time and place in a novice physician's career.

Those developing SP programs must concern themselves with issues of reliability, particularly in the training of SPs and the norming of their simulations, measures of consistency of scoring procedures, and so on. But once this is done, this study suggests that we might still expect moderate test score reliability in the communication domain. The variation and the instances of seeming contradiction across and within cases can and should be addressed as a technical measurement issue, but may also provide insight into students' level of development and have rich pedagogic value, providing specific entry to the multiple interacting dimensions of clinical performance. The richness of the SP narratives can be used to add a dimension to the checklist scores from Objective Structured Clinical Examinations in providing more explicit feedback to students. Standardized patients have a broad range of experience with a single problem, unlike a regular patient. They can compare across instances of the same event. They can compare the novice to the gold standard of the expert who trained them. They can speak clearly about strengths and weaknesses without fear of compromising their care. And they have an immensely larger sample of a student's clinical behavior than most attending physicians.

In addition, the variability in themes across the cases in the present study suggests a need to re-examine the way in which communication skills are taught so that both the general nature of these skills and their domain-specific application are explored by students. At present, most medical curricula teach communication skills as a set of generic skills with some attempt at advanced levels to make the psychosocial challenge more intense, e.g., giving bad news, talking with a run away teen, managing a hateful patient. This study suggests that more variables might be considered when increasing the complexity of communication challenges. How is communication affected when the diagnosis is unknown or more complex? A multi-institutional patient satisfaction study (Meredith and Wood, 1996) demonstrated that in real patients, the more serious the patient's condition, the more dissatisfaction the patient reported with the physicians' communication skills. Would a heart attack require a different communication approach than a diabetic patient?

More study of this emerging concept will be needed before we can claim with certainty that domain-specificity is a characteristic of communication skills.

References

Association of American Medical Colleges. (1999). *1999 Medical School Graduation Questionnaire Final School Report.* Washington, DC: AAMC.

Blue, A.V., Chessman, A.W., Gilbert, G.E. & Mainous III, A.G. (2000). Responding to patients' emotions: important for standardized patient satisfaction. *Family Medicine, 32*, 326-30.

Colliver, J.A., Willis, M.S., Robbs, R.S., Cohen, D.S. & Swartz, M.H. (1998). Assessment of empathy in a standardized-patient examination. *Teaching and Learning in Medicine, 10*, 8-11.

Colliver, J.A. & Swartz, M.H. (1997). Assessing clinical performance with standardized patients. *JAMA, 278*, 790-791.

Colliver, J.A., Swartz, M.H., Robbs, R.S. & Cohen, D.S. (1999). Relationship between clinical competence and interpersonal and communication skills in standardized-patient assessment. *Academic Medicine, 74*, 271-274.

Donnelly, M.B., Sloan, D., Plymale, M. & Schwartz, R. (2000). Assessment of residents' interpersonal skills by faculty proctors and standardized patients: a psychometric analysis. *Academic Medicine, 75* (October Supplement), S93-S95.

Hodges, B., Turnbull, J., Cohen, R., Bienenstock, A., & Norman, G. (1996). Evaluating communication skills in the objective structured clinical examination format: reliability and generalizability. *Medical Education, 30*, 38-43.

Holmboe, E.S. & Hawkins, R.E. (1998). Methods for evaluating the clinical competence of residents in internal medicine: a review. *Annals of Internal Medicine, 129*, 42-8.

Kurtz, S.M. & Silverman, J.D. (1996). The Calgary-Cambridge Referenced Observation Guides: an aid to defining the curriculum and organizing the teaching in communication training programmes. *Medical Education, 30*, 83-9

Meredith. P. & Wood, C. (1996). Aspects of patient satisfaction with communication in surgical care: confirming qualitative feedback through quantitative methods. *International Journal for Quality in Health Care, 8*, 253-64.

Miles, M.B. & Huberman, A.M. (1984). *Qualitative Data Analysis: A Sourcebook of New Methods.* Beverly Hills, CA: Sage Publications.

Rose, M. & Wilkerson, L. (in press). Widening the lens on standardized patient assessment: what the encounter can reveal about the development of clinical competence *Academic Medicine*, 76.

Spiro, H.S., Peschel, E. & St. James, D. (Eds). (1993). *Empathy and the Practice of Medicine: Beyond Pills and the Scalpel.* New Haven, CT: Yale University Press.

Vu, N.V. & Barrows H.S. (1994). Use of standardized patients in clinical assessments: recent developments and measurement findings. *Educational Researcher, 23* (3), 23-30.